MPRA
Munich Personal RePEc Archive

Gaming in a benchmarking environment. A non-parametric analysis of benchmarking in the water sector

De Witte, Kristof and Marques, Rui
University of Leuven (KUL), Technical University of Lisbon

12. April 2009

Online at http://mpra.ub.uni-muenchen.de/14679/
MPRA Paper No. 14679, posted 15. April 2009 / 09:42

Gaming in a benchmarking environment. A non-parametric analysis of benchmarking in the water sector.[1]

Kristof De Witte
University of Leuven (KUL)
Center for Economic Studies
Naamsestraat 69
3000 Leuven, Belgium
Tel: 0032 16 32 68 50
Kristof.dewitte@econ.kuleuven.be
www.econ.kuleuven.be/kristof.dewitte

Rui Cunha Marques
Technical University of Lisbon
Center for Urban and Regional Systems
Avenida Rovisco Pais
1049-001 Lisbon, Portugal
Tel. 00351 21 8418319
rcmar@civil.ist.utl.pt

April 12, 2009

Abstract

This paper discusses the use of benchmarking in general and its application to the drinking water sector. It systematizes the various classifications on performance measurement, discusses some of the pitfalls of benchmark studies and provides some examples of benchmarking in the water sector. After presenting in detail the institutional framework of the water sector of the Belgian region of Flanders (without benchmarking experiences), Wallonia (recently started a public benchmark) and the Netherlands (introduced already in 1997 a public benchmark), we non-parametrically measure the productivity gains by the use of a dynamic Malmquist index. The three regions, each at a different stage of the benchmarking circle, exhibit different performance trends. The 'carrot' and the 'stick' of benchmarking seem to offer an effective incentive to trigger performance. In addition, the Malmquist decompositions provide some evidence on the 'gaming' of the stakeholders by the water utilities.

Keywords: Benchmarking; gaming; Malmquist decomposition; regulation; water sector.

1. INTRODUCTION

Performance assessment encompasses the evaluation of efficiency (doing the things right) and effectiveness (doing the right things) of activities and processes and it is used for several ends and in different fields. If performance is assessed in the water sector, the utilities generally show a large potential performance gap. This can mainly be attributed to the presence of market failures, such as monopoly power, scale and scope economies, sunk costs and assets, services of general economic interest obligations and asymmetric information (Marques, 2005). In this context, performance assessment turns out to be an effective tool to guide managers, saving resources and improving the quality of service delivered.

[1] We would like to thank seminar participants at the Portuguese regulatory agency IRAR (Lisbon, Portugal).

As benchmarking and performance assessment are only light-handed regulatory tools, their importance is often undervalued. In this paper, we discuss the terminology, the classification and the main concerns that should be taken into account when benchmarking is employed. We apply the concepts and nomenclature to the water sector, although extensions to other sectors are generally straightforward.

Besides an extensive discussion on the concept of benchmarking, this paper contributes to the literature by an in depth analysis of the Dutch, Flemish and Walloon (the latter two are regions of Belgium) water sector. We especially emphasize the institutional framework and its benchmarking experiences. As the three regions are very similar in most of the relevant exogenous environmental characteristics (such as population density, geographical relief, weather conditions, expectations and demand of the customers, etc.), and as they have different experiments with benchmarking, it is very interesting and relevant to compare them from an institutional point of view. Indeed, whereas the Flemish region did not implement benchmarking or economic regulation, the Walloon sector started recently a public benchmark in the sense of 'sunshine regulation' (i.e. the comparison of utilities is used to 'embarrass' the least performing utilities and to put the best performing utilities into the limelight). Contrarily, the Dutch sector started a sunshine regulatory benchmark already in 1997.

As a third contribution to the literature, we estimate the performance of the water utilities in the different regions over time. Therefore, we apply a non-parametric (i.e. without *a priori* assumptions on the functional form of the production process) Malmquist index based on data envelopment analysis (DEA). This allows us to decompose productivity gains into gains arising from technical progress and gains from catch-up efficiency. To our best knowledge, this is a first similar decomposition for water utilities in the Low Countries.

Finally, we interpret and link the Malmquist decompositions to a 'gaming' framework (Jamasb *et al.*, 2003, 2004). Indeed, water utilities, as any economic agent, have their own agenda and act and react to incentives. The benchmark and the public debate provide incentives to the utilities. In turn, the utilities try to 'game' the system by changing their behavior

The paper unfolds as follows. The next section discusses the use and pitfalls of benchmarking. Section 3 and 4 describe, respectively, the institutional framework of the Dutch and Belgian water sector. Section 5 explains the methodology, while Section 6 interprets the results. Finally, Section 7 analyses the gaming aspects and Section 8 provides the concluding remarks.

2. THE CONCEPT OF BENCHMARKING

2.1 Performance assessment

Different performance measures can be employed to assess performance, such as (1) performance indicators, (2) performance indices and (3) performance levels (Alegre, 2007). Firstly, performance indicators are a quantified ratio expressing the way or intensity by which a given activity is accomplished. Including one productive factor in the numerator and one in the denominator, they are only a partial a-dimensional (in %) or intensive (in e.g. euro/m³) measure of productivity (sometimes even distorted). Besides other productive factors that influence the performance of a particular activity or process, there are explanatory factors that become key elements in the analysis of the scores obtained, although they are not intended to be a characteristic of the organization (e.g. population density).[2] Performance indicators are mainly used by operators and regulators for supervising the quality of service. Secondly, performance indices aggregate various performance variables in one value. Popular performance indices are Total Factor Productivity and Data Envelopment Analysis (DEA). Although in comparison to performance indicators, indices could better capture reality, they often encompass judgment values (e.g. the discretionary adoption of variables or weights) and work as a black box making it hard to understand how the inputs are transformed into outputs. Performance indices are often specially tailored for regulators, mainly if they adopt performance based regulation in the tariff setting process, such as price cap and/or yardstick competition methods. Finally, performance levels are related to qualitative measures associated with the quality of service provided. They are represented by discrete categories such as 'very good', 'good', 'fair' and 'poor', and are particularly useful for the customers to assess the service provided when the quantification of the measures is troublesome.

2.2 Classifications of benchmarking

Performance assessment is not synonymous with benchmarking, even though the terms are often mixed up. Benchmarking is a narrower concept as it also encompasses the comparison with other peers and standards. Camp (1989) refers to benchmarking as "*the search for industry best practices that lead to superior performance*".

Traditionally in the water sector, benchmarking methods are classified into metric benchmarking and process benchmarking (Kingdom *et al*., 1996). The former refers to the comparative and quantitative process that enables the operators to keep track of their performance through time and compare it with other (similar) operators. The latter identifies, in a first phase, the aspects to be improved through a step-by-step procedure (process mapping) and, in a second phase, compares them with the best practices from other operators. Metric benchmarking intends to answer the question about *what* to improve, whereas process benchmarking aims at answering the question of *how* to improve. In spite of other useful classifications of benchmarking, depending on the technique used (e.g. total or partial methods, non-parametric *versus* parametric techniques) or the aim (e.g. regulatory or non-regulatory), we can look at these tools concerning their end users.

Metric benchmarking determines the current performance levels and establishes appropriate reference values. However, process benchmarking goes one step beyond as it discriminates the existing processes in a given organization, compares them with other organizations with exceptional performance in similar processes (best in class) and defines changes regarding each process to achieve or overcome the performance aims set.[3]

---

[2] For example, a performance indicator that intends to analyze the productivity of staff can be captured by the number of employees per customer, per mains length or per volume of distributed water. On the one hand, different indicators can capture the performance of the same variable. On the other hand, two equal scores do not have the same evaluation if the network age differs.

[3] This type of benchmarking, known as "*Xerox-style* benchmarking", was developed by the Xerox Corporation in the 1970s in order to compete with the Japanese enterprises (Kingdom, 1998).

Clearly, there are two schools of benchmarking practitioners, i.e. those focused on economical issues and those motivated by operational aspects. On the one hand, the economical issues are of interest to regulators, government agencies and supra-national institutions which mostly use metric benchmarking, particularly performance indices or performance indicators. On the other hand, operational issues are, among others, addressed by water utilities that are using process benchmarking, mainly by performance indicators.

In Figure 1, we schematically present a metric benchmark. Firstly, for the evaluated activity, it has to be decided on the performance measure and the measurement criteria (e.g. to measure the water losses the non-revenue water volume by pipe extension and per hour can be used as criterion or, with the same aim, the non-revenue water by customer and per day). Secondly, after computing the measure, the explanatory values that may influence the performance are examined. They are divided into controllable and non-controllable, depending on the capability of the utilities to tackle them in the long run. For example, managers can influence the age of assets in the long but not in the short run. In its turn, the age of assets have considerable consequences in leakage. Therefore, performance indicators concerning leakage should take this explanatory factor into account. In addition to leakage, for instance, the type of soil is always non-controllable. Besides explanatory factors, we should establish optimal or reference values for the performance measure. If the best practices are not available or are far from the current performance, the reference values can be compared with scores from past periods. The comparison with reference values is crucial in the analysis. Indeed, the utilities should use these reference values as objectives to accomplish in the following periods.

When using performance indicators, despite the different arrangements defined in the literature (e.g. by the International Water Association (IWA), the American Water Works Association (AWWA) or the World Bank systems), one should note that there are no performance indicators systems *a la carte*.[4] A performance indicators set should be defined according to the needs of the evaluated entity, trying to use the most adequate language and terminology.

Figure 2 shows one of the first cases of process benchmarking in the water sector. It corresponds to the water leakage reduction plan conducted by the Philadelphia Water Department. Although neither regulators nor other government agencies do employ process benchmarking, they have all the motivation to encourage this kind of exercises in the water sector because, as a rule, they contribute for performance improvement (Parena *et al*., 2002). Process benchmarking produces good results, showing "how to improve" and not only "what to improve".

[4] See the IB-NET tool kit (www.ib-net.org/), Alegre *et al*. (2006) and Lafferty and William (2005).

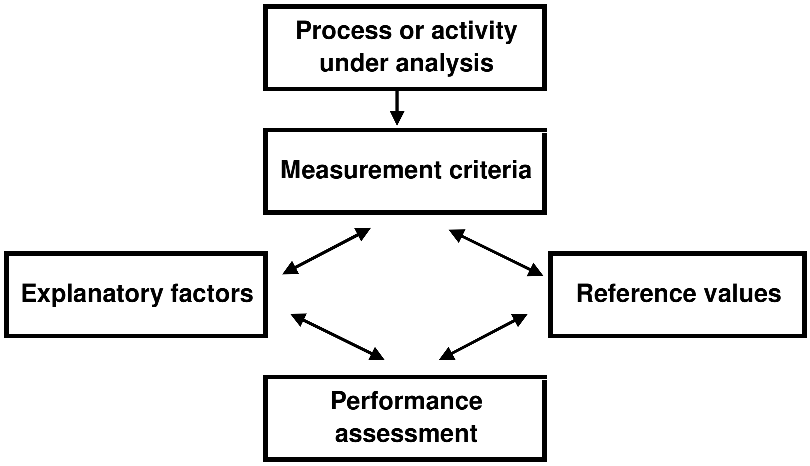

Figure 1 – Application scheme of metric benchmarking

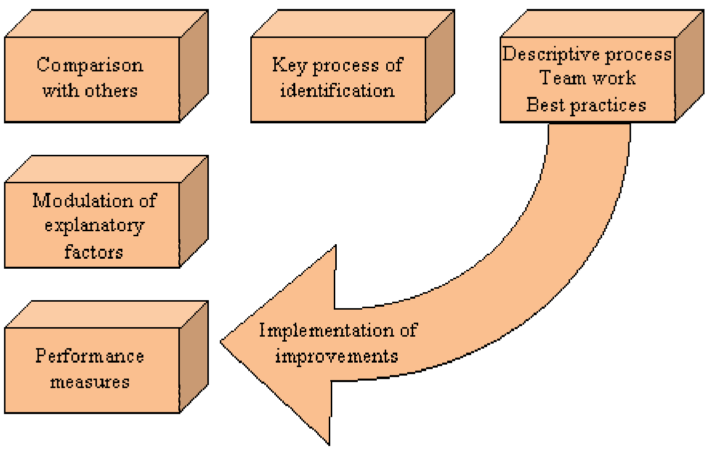

Figure 2 – Application scheme of process benchmarking (Freeman *et al*., 1998)

2.3 Pitfalls of benchmarking

The application of benchmarking leads to a better knowledge of the other operators. It allows the operators to reduce the performance gap to achieve the best practices. However, in applying benchmarking some cautions should be taken (see Marques and Witte, 2008). Firstly, the organizations or activities which are subject to benchmarking must be comparable (like with like). Although heterogeneity can be modeled and included in the analysis, a very big difference will bias the analysis (e.g. a water utility of 1,000,000 inhabitants should not be compared with a utility of 1,000 inhabitants). Secondly, to avoid a simplistic analysis, all variables should be weighted. Thirdly, benchmarking requires patience as the process is not immediate and the results are not promptly tangible. It is necessary to look for a consensus among stakeholders to assure continuity through time and to define medium and long term goals so that the system of performance measurement may be successful and produce the desired results. Benchmarking application, as well as performance measurement in general, should not be an aim in itself. Fourthly, in benchmarking some results are expected to be obtained. Performance should be objectively evaluated by displaying the areas where improvement is needed and by identifying other organizations whose processes show a higher performance with the aim of adopting them and testing whether the improvement programs are successful. Fifthly, the generated transparency allows for a better understanding of performance as benchmarking allows for insights in the way the utility works. However, as a drawback, the probability of a *status quo* in the organization being affected is large.

The scope of the benchmark should comprise the organization as a whole, focusing on the objectives and the customers. Besides, benchmarking provides more questions than answers. The evaluation should focus on and provide evidence of the aspects that the managers are able to control. Ranking by itself should not be an end to the process of evaluation. On the contrary, it should be a tool for change towards excellence. Finally, it should be highlighted that benchmarking is a cyclic and permanent process. The identification, understanding and implementation of best practices should not be seen as an occasional and solitary task, but as an activity that is entangled in the organizational culture where there is a need for continuous improvement. Nevertheless, one should be aware that the process of improvement is continuous and time consuming.

2.4 Benchmarking experiences in the water sector

Benchmarking is being applied worldwide in the water sector. Governments and professional associations in each country are promoting benchmarking exercises as a leverage on performance improvement. The water utilities are also more and more interested in the use of benchmarking. Even in countries without regulators, as the majority in Europe, benchmarking experiences have been in progress with success. A new regulatory paradigm based on yardstick competition, and therefore on benchmarking, has been proposed for the water sector (Marques and De Witte, 2008).[5] For example, in Europe there are experiences of benchmarking by the regulators as in Portugal (Institute for the Regulation of Water and Waste - IRAR), England and Wales (Water Services Regulation Authority (Ofwat)] and Belgium (in Wallonia by the Committee for Water Control), by the association of water utilities as in Sweden [The Swedish Water & Wastewater Association (SWWA)], Denmark [Danish Water and Waste Water Association (DANVA)] or Holland (Association of Dutch Water Utilities - VEWIN) and by the water utilities consortia like 6-cities in Scandinavia or the Water Utility Management Plan (WUMP) 2050 in Sweden. Other initiatives are being developed in Germany, Norway, France, Slovenia, Slovakia and Czech Republic. Outside Europe the picture is not different. Besides, the supra-continent projects like the ones of IWA, Word Bank or World Health Organization, Australia, Brazil, Argentina, the USA and Canada are some of the good examples of benchmarking application. In the next section we analyze in detail the experiences of the Low Countries. The development of the institutional framework of the water sectors in Belgium (in particular its regions Flanders and Wallonia) and in The Netherlands has been related to their performance assessment. The water utilities, intending to avoid reforms, 'game' their respective governments and stakeholders looking for the maintenance of the *status quo* and the *quiet life*. This gaming is discussed in Section 7

[5] Yardstick competition always implies the application of benchmarking but it is not a regulatory method. There are two types of yardstick competition in the water sector, respectively the 'price yardstick competition' and 'sunshine regulation' (Marques, 2006). The latter entails the comparison and public discussion of performance of the water utilities.

after we carefully described the institutional framework in the drinking water sectors of Flanders, Wallonia and the Netherlands.

3. THE DUTCH WATER SECTOR

As a water rich country, since the middle ages the Netherlands had to learn how to manage their water resources in order to survive and to recover land on the sea. This involvement of all citizens in the fight against water is still reflected in the current regulatory and institutional framework. For example, the Dutch citizens are heavily involved in the water production, distribution and preservation process by the democratically elected ‘water boards’, which are an inheritance of the medieval fight against water. Although the Dutch Ministry of Economic Affaires launched in 1997 a report on the privatization of water services, only six years later is the drinking water sector protected as fully public domain. In this section, we describe the institutional and regulatory changes in the Dutch drinking water sector which lay the foundation of this policy shift.

3.1 The institutional framework

The legislative powers in the Dutch drinking water sector are, in the unitary decentralized state, divided over the State, the provinces, the municipalities and the water boards. The latter is a special governmental agency responsible for the water management. At the State level, the Ministry of Housing, Spatial Planning and Environment (VROM) and the Ministry of Transport, Public Works and Water Management (V&W) are ultimately responsible for the largely decentralized water sector. The former concerns the drinking water sector, while the latter deals with the sewage sector and surface water quality. Even though the provincial authorities can outline their own policy plans, they mainly act as a coordination device between the State, the municipalities and the water boards (in both directions). In addition, they are shareholders in the water utilities and in the performance of this role they supervise the drinking water tariffs (together with the municipalities). The municipalities are also responsible for the collection of sewage. The most typical Dutch drinking water institutions are the water boards which are responsible for the water management (including e.g. the dunes, dikes and levees), water infrastructure, water quality and the treatment of sewage. In 2007, there were 27 water boards consisting of democratically elected representatives of four relevant categories of stakeholders: residents of the area, landowners, company owners and real estate owners. In recent years, the turnout percentage at elections was rather low (25%). The continuous monitoring of the water quality by both the water boards and the other government levels makes it relatively easy for the Dutch water utilities to comply with the strict European directives. The drinking water companies are responsible for the delivery of water in their legal monopolistic provision area.

The utilities faced two merger waves in the past. In the first merger wave, around 1920, the companies searched for a scale increase to comply with new environmental legislation. The second merger wave, around the late 1990s, was originated from the search for scale economies, improved corporate governance, stricter environmental requirements and the pressure of the national and provincial authorities to have one supplier per province (De Witte and Dijkgraaf, 2007). While there were 111

utilities in 1975, 32 companies in 1997, only 10 companies remained in 2007 (although the utilities can still be labeled as medium sized companies in comparison with, e.g., some English companies). It is expected that in the near future only 3 to 6 companies will prevail.
Since 2003, the national government reserved the sector as a public domain, which implies a moratorium on private investments. Therefore, only the provincial and municipal authorities can be the shareholders of the utilities which are organized as limited companies under company law (i.e. Public Limited Company, PLC). The drinking water act (temporarily) settled the 1997-debate on the privatization of water services as the water companies are allocated to an exclusive provision area (although large customers with more than 100,000 m³ per year can choose their supplier). Both the water boards and the drinking water utilities are coordinated in sector organizations, respectively, the Association of Water Boards (Unie van Waterschappen) and VEWIN (Vereniging van Waterbedrijven). Particularly the latter is of importance for the remaining of this section, as it organizes the benchmark since 1997 (Van Dijk *et al.*, 2007).

3.2 The benchmark

In 1997, the Dutch Ministry of Economic Affairs launched a project to deregulate markets with monopolistic power, among which the drinking water sector. In an influential study, Dijkgraaf *et al.* (1997) compared the possibilities for different degrees of competition in the market. As the sector was not influenced by incentives to monitor and improve the efficiency, it was expected that increased market processes would enhance efficiency. To avoid competition in or on the market, and to circumvent the liberalization and privatization idea, VEWIN started in 1997 on a *voluntary* basis a benchmarking process (by the help of an external consulting company to carry out the benchmark and by accountants to check the values). By the use of performance indicators for Water Quality, Services, Environment and Finances, on both the company level, process level (e.g. production, distribution, sales and general management) and sub-process level (e.g. cost per meter pipe), the benchmark was expected to generate information and to provide transparency in the sector. Started as a voluntary self-regulation project, the benchmark is made compulsory by law (2005). The utilities receive yearly a detailed internal report and every three years an external report is published. The latter is closely followed by the national media and the Dutch citizens. Before disclosing new variables to the general public, drinking water utilities report the variables internally in the two proceeding years. Besides the internal carrots (managers negotiate contracts with bonuses depending on the benchmark results) and external rewards (the benchmark outcomes are closely monitored by the public opinion which are the voters of the municipal delegates), the government does not attach bonuses nor sanctions on the benchmark outcomes (i.e. self-regulation by the sector).
The benchmark proved to be highly effective. Dijkgraaf *et al.* (2007) and VEWIN (2007a) estimated the efficiency improvement of the benchmark at 23% since 1997. In addition, also service levels, investments and quality significantly improved. Although the efficiency in the sector drastically improved, nominal prices did not decrease, so that profits of the drinking water companies expanded (for a discussion on the profits and profit drivers, De Witte and Saal, 2008). The accumulation of profits results in a high solvability of the sector (on average 30%) and high returns on net assets (for

the three largest companies in 2006, 12.5%, while e.g. in England and Wales the return is 6.5% before tax).
Furthermore, the current benchmark, based on performance indicators is not perfect. Besides figures for profits, the benchmark should account for exogenous characteristics which are beyond the scope of the firm's managers. The quality of the drinking water provision area differs, so that uncorrected performance indicators are partially incomparable. Also small differences in accounting rules and in employed definitions make a fair comparison intricate.
The decreasing number of drinking water utilities diminishes the effectiveness of benchmark as the number of comparable units falls. Therefore, VEWIN started, at the end of 2005, an international benchmark with some Scandinavian utilities, and by the end of 2007 its scope was enlarged to English and German utilities (VEWIN, 2007b)

4. THE BELGIAN WATER SECTOR

As a federal country, the Belgian water industry has been regionalized since 1980. This shift in responsibilities from the national state level to the regional state level, significantly altered the existing homogeneity in the sector as the Flemish, Walloon and Brussels region are characterized by a large structural, cultural and geographical diversity (Aubin and Varane, 2007). In addition, the Flemish and Brussels region have few water resources of their own, so that they are, respectively, about 25% and 97% dependent on Wallonia. Thanks to strict European legislation and the increasing awareness of the customers, the previously little regulated sector is changing. In this section, we discuss the regulatory and legislative framework of the Flemish, Walloon and Brussels water sectors and briefly explore the Walloon benchmark.

4.1 Institutional framework

Since the establishment of Belgium in 1830, the municipalities hold the responsibility to provide drinking water to their citizens. Initially, the sector was dominated by municipal services (on the payroll of the municipality) and municipal suppliers to supply drinking water. However, to respond to the capital intensity of the sector and the continuously stricter legislation, the municipalities organized themselves into inter-municipal utilities. Nowadays, the Flemish region counts only 6 municipal suppliers and 3 inter-municipal utilities. Although the Walloon government insists by the 'code de l'eau' on a rationalization of the number of drinking water companies, it still comprises some municipal services and municipal suppliers and 13 inter-municipal utilities. If the municipalities refrained to supply drinking water, the national company (NMW) would take over this responsibility. Due to the regionalization, in Flanders the NMW was converted into the 'Vlaamse Maatschappij voor Watervoorziening' (VMW), while in Wallonia the 'Société Wallonne des Eaux' (SWDE) was established. The former provides 40% of the Flemish population, while the latter counts half of the Walloon connections. The Brussels drinking water is provided (wholesale segment) by Vivaqua (an inter-municipal utility) and distributed by another inter-municipal utility (Belgaqua, 2007; Serv, 2007).

In Flanders, the integrated water system passes the responsibility of sewage treatment to the drinking water suppliers. Therefore, the drinking water utilities agreed with the monopolistic Aquafin, an initial public-private partnership to treat the sewage but since 2006 it became fully public. Though the responsibility for sewage shifted from the municipalities to the drinking water suppliers, the former can still decide on who should execute these tasks in their territory: either themselves or the drinking water utility. In Wallonia, the water treatment is not the exclusive power of the 'Société publique de gestion de l'eau' (SPGE), the Walloon equivalent of Aquafin. Some inter-municipal utilities are providing both drinking water and treating the sewage themselves. In Brussels, the 'The Brussels Water Company' (BMWB) is responsible along with Vivaqua for the sewage treatment (Serv, 2007). Table 1 systematizes the Belgian market structure in 2007.

Table 1 – Belgian market structure

	Direct public		**Delegated public**		**Delegated**		**Direct public**	
	W	S[1]	W	S[2]	W	S	W	S
Population	10%	100%	85%	100%	5%	0%	0%	0%
Operators	47		47		10			
Trend	↓		↑		↑			

W – Drinking water supply; S – Sewage; [1] – Collection; [2] - Treatment

The regionalization of the sector in 1980 transferred the legislative authority from the national government to the regional authorities. At the national level, Belgaqua's main task is to represent the utilities at the national and supra-national forums. Although the regions did not establish a regulatory body *stricto sensus*, some important regulatory functions are held by the sector organizations. However, the Flemish government agreed to establish by July 2003 a regulatory entity whose aim was to survey, evaluate and provide advice on the transparency, the public service obligations, the investments, cost structures and yardstick competition. In contrast to other countries and regions, only few initiatives to promote efficiency and effectiveness in the Flemish drinking water sector have been developed. Despite the fact that most of the utilities provide information on the tariffs on their website, a full overview of the different companies is not available. In addition, none of the companies provide insight in the cost determining factors. In 2002, the Flemish sector organization SVW organized a benchmarking exercise after the Dutch example. Each participating company received a document with its own performances relative to a reference group. However, these results have never been released to the general public. From the three Belgian regions, the Walloon water sector is the most regulated one. The Commission for Water Control (CWC), integrated in the public administration, controls and supervises the tariff systems (although the final responsibility belongs to the federal Ministry of Economic Affaires), observes the legal obligations of the operators and implements performance indicators in order to increase the efficiency in the sector. In Brussels, the sector is supervised by the public administration in 'Environment Brussels', though no regulatory body is established (Aquawal, 2006).

The tariff system differs significantly among the three regions and even within operators, although all prices are supervised by a federal pricing commission (part of the federal Ministry of Economic Affaires) which determines the maximum prices. This can be explained by difference in the network length, the degree of urbanization, the type of clients (i.e. size of the customers), the relative abundance of water and the source of water extraction.

A detailed request should be submitted at least 60 days before the scheduled price increase. The following elements should be contained in the proposal: (1) the refusals and approvals of previous price increases; (2) the date of the previous price increase; (3) the full or partial price increase; (4) a detailed description of the tariff system; (5) the evolution in the volume of water use; (6) a comparative study with the results of the last three fiscal years; and finally (7) a detailed analysis of the costs and the revenues.

In the Flemish region, the government determined to provide each domestic customer 15m³ drinking water for free (paid by increasing the price of the remaining volumes). This initiative to help disadvantaged people is seen by the sector as a powerful sign of the public aspect of the sector. The remaining consumption is billed by a fixed component and a variable component. The latter part is in general uniform or regressive (which favors medium and large sized customers), and only occasionally progressive (only for three operators). In general, the drinking water utilities charge the same fee both for small and large sized customers, although the very large customers can negotiate water prices individually. In addition to the 15m³ free drinking water, only two inter-municipal utilities give a reduction for low income customers. The sewage tariff consists of two components which are both collected by the drinking water utilities: the sewage treatment and sewage collection. The former concerns the fee for Aquafin (in 2006, 0.6798 euro per m³ without VAT of 6%), while the latter encompasses the costs of the municipality to collect sewage (autonomous determined fee by the municipality, with a maximum of 1.5 times the treatment costs). The law foresees for large sized customers a reduction in the sewage costs. We present the tariff system in 2006 in Table 2.

In the Walloon region, by the 'code de l'eau', the government tries to make the water price uniform in the same water basin. Since 2005, the water price consists of a fixed and a variable component, regardless of the nature of the customers (domestic or non domestic). Both parts should reflect the 'truth-cost' of water and sewage provision. The variable component comprises four parts. Firstly, a progressive real distribution cost (CVD) which should be the same for all customers in the water basin. The CVD is yearly set by the federal pricing commission after a screening by the CWC. Secondly, a contribution of 0.0125 euro/m³ is invoiced for a social fund for disadvantaged people. Thirdly, a uniform fee for the Walloon region (determined by the SPGE) is levied to finance the water treatment (CVA). Finally, as in Flanders, the VAT amounts to 6% (Serv, 2007). The tariff system is presented schematically in Table 2.

In Brussels, the tariff system for domestic customers consists of a fixed component and a progressive variable component, which is increased by a contribution of 0.01 euro/m³ for a social fund to allow reimbursement of the most disadvantaged families. The sewage contribution depends on local conditions (municipalities or regional system) and includes a fixed portion and two tiers variable to the consumption.

Table 2 – Tariff system for domestic customers in Belgium

Region	**Fixed component (€)**	**Variable component (€ / m³)**		**Social fund (€ / m³)**	**Wastewater tariff (€ / m³)**	
Flanders	39 – 57	0-15 m^3	0	15 m3 free	Aquafin	Municipal

		>15 m^3	1,881-2,931		0,7730	variable
Wallonia	20CVD + 30CVA	0-30 m^3 30-5000 m^3 5000-25000 m^3 >25000 m^3	0,5CVD CVD+CVA 0,9 CVD+CVA 0,7CVD+CVA	0,0125	Together with the water sold	
Brussels	11,9 - 23,80	0-15m³ 15-30m³ 30-60m³ >60 m³	0,80 1,39 2,06 3,06	0,01	IBDE Regional 0,3275	IBDE Municipal 0,43

Public service obligations are not settled by a federal law. Each of the three regions, taking into account the European Directives (e.g. water quality standards or abstraction protection), defined a set of minimum conditions for the drinking water utilities. In Flanders, besides providing 15m³ of free drinking water, the drinking water utilities defined a set of minimum obligations for the customers. These standards foresee financial compensation (around 25 euro) if the targets are not reached.
In Wallonia, the ‘code de l’eau’ established a social fund to subsidize less favored customers and allowed the CWC to develop a set of 13 indicators to measure the legal obligations of the utilities. These include the access to information, the respect for privacy of customers, the identification of workers, the time to answer complaints and requests, the quality of service, the clearness of the invoice and the easiness for payment (Aquawal, 2006).
In Brussels, the main operator (IBDE) defined standard measures, aimed at protecting the interests of customers, but no financial contribution is provided if they are not respected.

4.2 The Walloon benchmark

Before the benchmarking project, which started in 2006 with first results expected by the end of 2007, the Walloon drinking water sector worked in a self-regulation system where public service obligations were not supervised, the quality of service was not controlled and sometimes not even known. The sector association, Aquawal, enhanced the standardization and dissemination of best practices and the disclosure of statistical information. To measure the performance of the utilities, the CWC developed, besides the public service obligations (see supra), a set of 15 performance indicators. These benchmark measures, calculated and provided by the operators, are divided into six classes: quality of service, abstraction protection, management and sustainability, pricing and management, coverage and solidarity and satisfaction of customers and communication. By these indicators, CWC wants to estimate and compare the evolution of the quality of service among operators (i.e. a sunshine regulation), and tries to create and stimulate the competition between the utilities (i.e. a yardstick competition or competition by comparison). In addition, the performance indicators and their target values allow for the implementation of a management system by objectives. The inclusion of context variables helps to interpret the performance indicators. Figure 3 presents systematically the scheme of CWC service regulation.

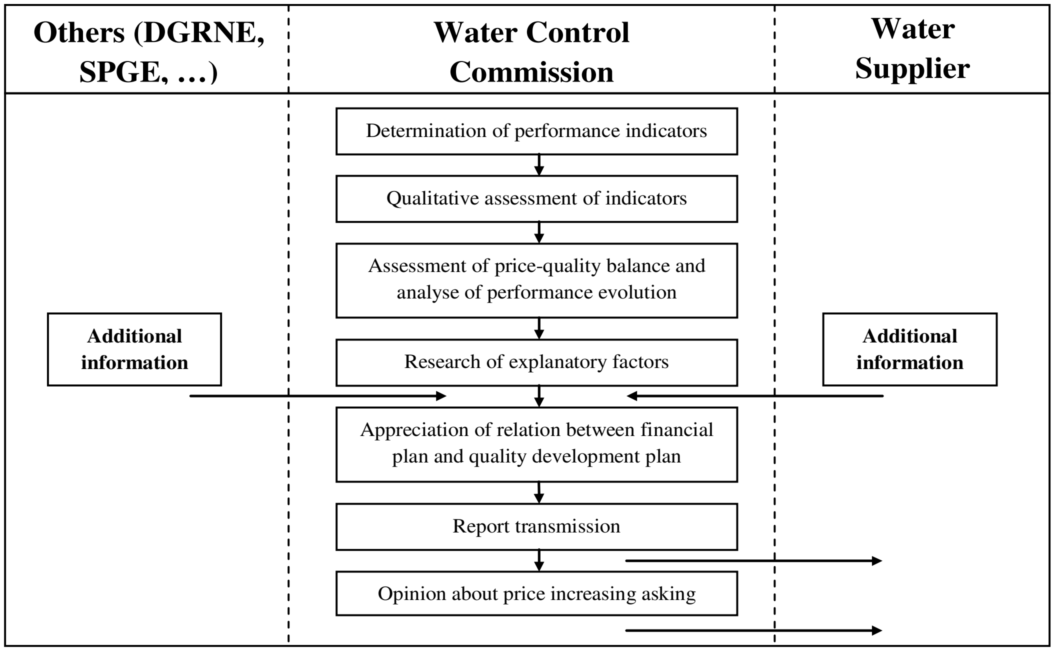

Figure 3 - Scheme application of performance indicators (Aquawal, 2006, p. 50)

5. MEASURING EFFICIENCY OVER TIME

Although the Netherlands, Flanders and Wallonia share many exogenous factors (in fact, they are working in a very similar environment with a high population density, water rich underground, comparable weather conditions, analogous expectations from the customers, similar geographical relief, etc.), the major difference lies in the stage of the benchmarking cycle. Whereas Flanders has not implemented a public benchmark yet, the Walloon utilities have started to benchmark only recently and the Dutch utilities have been self-regulated by benchmarking since 1997. This observation makes a comparison through time especially relevant. In particular we verify to which extent the resources are used to produce outputs and relate this efficiency score to the stage of the benchmarking project.
To estimate the efficiency non-parametrically (i.e. without assuming any *a priori* assumption on the production frontier), we start from a best practice frontier. Consider therefore the two-dimensional Figure 4. We present the inputs x ($x \in \mathbb{R}_+^p$) on the horizontal axis and the outputs y ($y \in \mathbb{R}_+^q$) on the vertical axis. In this two dimensional example, $p = q = 1$. The observations (also labeled as Decision Making Units, DMUs) are transforming the inputs into outputs with a different success. Consider observation DMU_1 which uses x inputs to produce y outputs. Some observations are able to produce at least the same amount of outputs, with at most the same quantity of inputs (e.g. DMU_2 to DMU_5), so that DMU_1 is dominated by these units. The undominated observations are considered as best practices.

The non-parametric Data Envelopment Analysis (DEA) model (Charnes *et al.,* 1978) estimates efficiency of an observation relative to a convex hull around these best practice observations. In particular, the DEA model assumes free disposability of the inputs and outputs: $\forall(x, y) \in \Psi$, if $\tilde{x} \geq x$ and $\tilde{y} \leq y$ then $(\tilde{x}, \tilde{y}) \in \Psi$, or in words: if a particular input-output combination (x, y) is feasible (i.e. it consists of the production set $\Psi = \left\{(x, y) \mid x \in \mathbb{R}_+^p, y \in \mathbb{R}_+^q, (x, y) \text{ is feasible}\right\}$), it should also be possible to produce *y* by more inputs and to produce by the input level *x* also less outputs. The corresponding best practice production set is defined as the set of undominated observations (undominated in both the input and output dimension):

$$\Psi = \left\{(x, y) \in \mathbb{R}_+^{p+q} \mid x \geq x_i , y \leq y_i , i = 1,...,n\right\}. \qquad (1)$$

Additionally to the free disposability assumption, DEA assumes a convex shape of the frontier: if $(x_1 , y_1), (x_2 , y_2) \in \Psi$, then $\forall \alpha \in [0,1]$: $(x, y) = \alpha(x_1 , y_1) + (1-\alpha)(x_2 , y_2) \in \Psi$. As such, the corresponding best practice production set is defined as a convex hull of the undominated input-output combinations:

$$\Psi_{DEA} = \left\{(x, y) \in \mathbb{R}_+^{p+q} \mid x \geq \sum_{i=1}^{n} \gamma_i x_i , y \leq \sum_{i=1}^{n} \gamma_i y_i , \text{for } (\gamma_1,...,\gamma_n), \right.$$

$$\left. s.t. \sum_{i=1}^{n} \gamma_i = 1, \gamma_i \geq 0, i = 1,...,n\right\}. \qquad (2)$$

The assumption that $\sum_{i=1}^{n} \gamma_i = 1$ assures that the best practice frontier is a convex hull around the observed data, and as such, implies variable returns to scale (VRS; Daraio and Simar (2007) prove VRS to be consistent whatever underlying model, and De Witte and Marques (2007b) claim its appropriateness for the water sector).

As suggested by Farrell (1957), we can measure the relative efficiency of entities as the distance to the best practice frontier. The input-oriented approach (or horizontal inefficiency) searches for the minimal amount of inputs which an observation should use to produce a given amount of outputs if it would be as efficient as its best practice. Algebraically, this corresponds to $\partial(x, y) = \inf\left\{\partial \mid (\partial x, y) \in \Psi\right\}$ where $\partial = 1$ denotes efficiency and $\partial < 1$ inefficiency. Graphically, as presented in Figure 4, this corresponds to the relative distance AB/AC. The output-oriented approach (or vertical inefficiency) measures the maximal amount of outputs which could be produced with a given level of inputs if the observation is working as efficient as its best practice. This is algebraically measured as $\lambda(x, y) = \sup\left\{\lambda \mid (x, \lambda y) \in \Psi\right\}$ where $\lambda = 1$ denotes efficiency and $\lambda > 1$ inefficiency. This is reflected by the distance AE/ED. In this paper, we will focus on the input-orientation as due to demand side restrictions the outputs of a water utility are fairly exogenous.

A major criticism on the non-parametric DEA models lies in their deterministic nature and inability of statistical inference. Recently, Cazals *et al.* (2002) developed a more robust approach to account for these criticisms. Basically, the procedure consists of repeatedly and with replacement drawing *m* observations from the data set for which $y \leq y_i$. As such, it constructs a partial frontier with *m* units (in contrast to the full frontier with all observations) relative to which we measure the efficiency by the input-oriented DEA problem. The procedure is repeated *R* times, after which the results are averaged so that we obtain the robust input-oriented DEA efficiency score. The major advantage of the order-*m* approach lies in its possibility to draw statistical inference and in its robustness to outliers. The robust order-*m* approach of Cazals *et al.* (2002) is incorporated in our input-oriented DEA model.

The above described DEA model is only valid for cross-section samples. In more interesting applications, we evaluate efficiency in a panel data set by the use of a Malmquist index (Caves *et al.,* 1982). The Malmquist index allows us to measure the change in efficiency over time and to decompose the latter into (1) a frontier shift, which reflects the change of the best practice frontier, and (2) a catching-up effect, which estimates whether the entities are closer to the best practice frontier. The Malmquist index is basically a decomposition of the robust input-oriented DEA model over two time periods. As such, it compares the shift of the frontier Ψ between periods *t* and *t+1* (i.e. the evolution in efficiency of the best practices) and the efficiency of the evaluated observation relative to Ψ^t and Ψ^{t+1}. In the two dimensional Figure 5, we present a frontier shift between period *t* and *t+1*. As in period *t+1* the observations can produce more outputs with a given level of inputs, or vice versa, as in period *t+1* the observations need less inputs to produce the same quantity of outputs, the sector experienced a technological progress (i.e. $\Psi^t \subset \Psi^{t+1}$). However, for the evaluated observation DMU_1, the best practices (which constitute the frontier) improved their performances faster than observation DMU_1. Therefore, the observation experiences a negative catching-up effect between the two periods.

Algebraically, we measure the Malmquist index, using period *t* as benchmark technology, as (for a comprehensive description, Fried *et al.*, 2008):

$$M^t(x^t, y^t, x^{t+1}, y^{t+1}) = \frac{D^t(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)} \tag{3}$$

where $D^t(x^{t+1}, y^{t+1})$ denotes the efficiency of observation (x^{t+1}, y^{t+1}) with the reference technology of period *t* (i.e. compare observation (x^{t+1}, y^{t+1}) with observations in period *t*). An index of M^t <,=,> 1 denotes, respectively, a productivity growth, stagnation or decline between period *t* and *t+1*. However, to avoid the arbitrariness of choosing one or another benchmark period (indeed, selecting period *t+1* as benchmark period gives other results), the Malmquist index is conventionally defined as the geometric mean of the two:

$$M(x^t, y^t, x^{t+1}, y^{t+1}) = \left[\frac{D^t(x^{t+1}, y^{t+1})}{D^t(x^t, y^t)} \times \frac{D^{t+1}(x^{t+1}, y^{t+1})}{D^{t+1}(x^t, y^t)}\right]^{1/2} \tag{4}$$

where $M <,=,> 1$ denotes, respectively, a productivity growth, stagnation or decline between periods *t* and *t+1*.

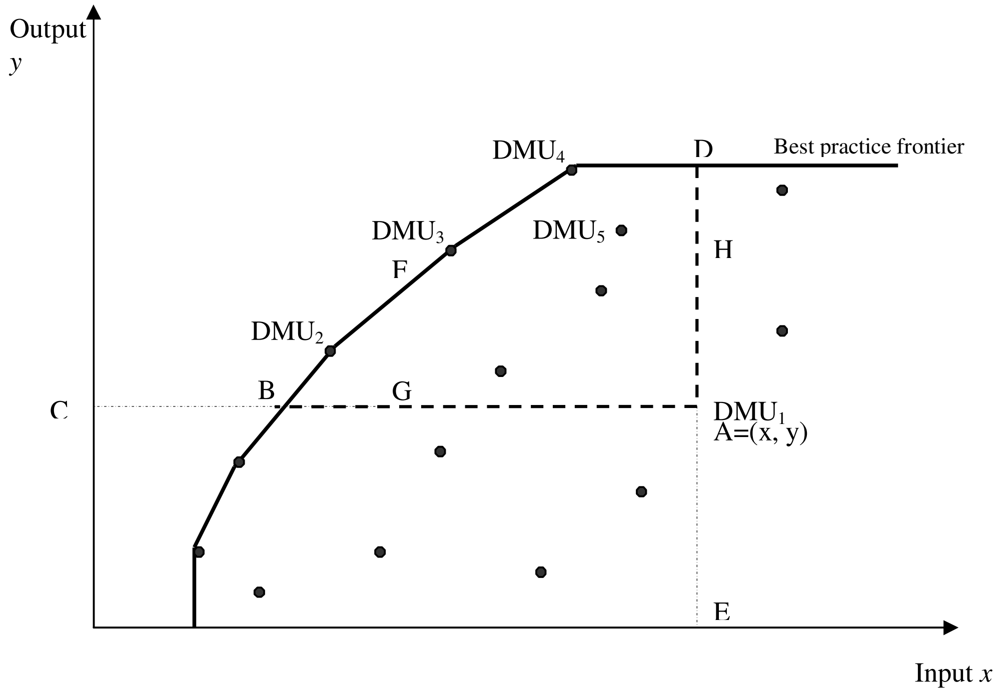

Figure 4 - The DEA approach

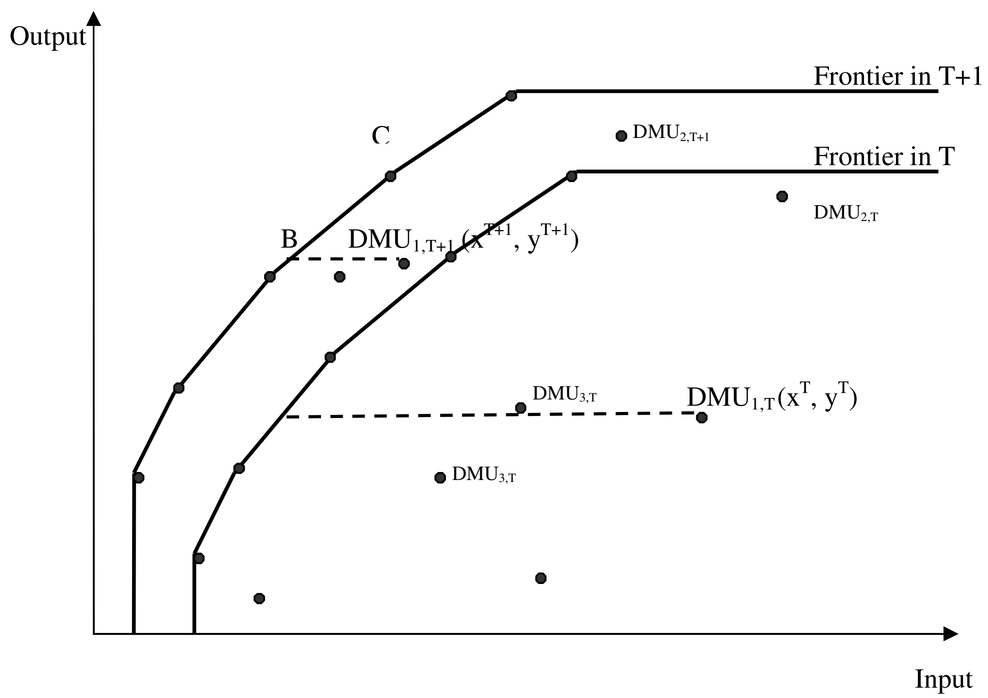

Figure 5 - Efficiency measurement over time

6. RESULTS OF THE MALMQUIST INDEX

To measure the frontier shift and efficiency change in the Belgian and Dutch drinking water sector, we apply the robust (Cazals *et al.,* 2002) input-oriented DEA model with its Malmquist decomposition in the panel data set. The data are obtained from the sector organizations Belgaqua and VEWIN and from the annual accounts of the drinking water companies. As a result of the mergers in the Dutch drinking water sector, the number of utilities gradually decreased over the last years. To avoid the decreasing number of utilities in the sample (which could intricate the analysis as a smaller number of observations reduces the consistency of the DEA estimates) and to avoid sample size bias (indeed, an efficiency evaluation in a smaller sample increases the average and individual efficiency estimates), similarly as in De Witte and Dijkgraaf (2007), we decompose the variables of the 13 merged companies in 2005 to the 20 companies in 1995 by extrapolating the change in the variables in the merged utility to the sub-utility (although we use only data from 1999-2005). For an extensive discussion on how each variable of the merged utility is decomposed to the variables of the sub-utility, we refer to De Witte and Dijkgraaf (2007). Summary statistics are presented in Table 3. In spite of having experimented with several input variables (e.g. total costs, wage base and capital base), we selected two input and two output variables which are consensual in the literature (De Witte and Marques, 2007b). In particular, we use labor and capital as inputs. These are, respectively, proxied by the number of employees (in FTE) and length of the mains network (in km). The output variables measure the number of connections and the total volume of delivered drinking water. As this corresponds to the fixed tariff per meter and the variable tariff per m³, it captures the revenues of the drinking water companies. To estimate the efficiency shifts over time, we collected data from 1999, 2001, 2003 and 2005. We selected 1999 as the first year of analysis as this corresponds to the presentation of the first public benchmark study in the Netherlands. In addition, De Witte and Dijkgraaf (2007) show that the effect of the benchmark is only noticeable from 1999 on. The data set consists of 27 Walloon, 5 Flemish and 20 Dutch utilities.[6]

The results of the Malmquist decomposition are presented in Table 4. Values larger than 1 denote a relative productivity growth, negative values indicate a decline and values equal to 1 point to a steady state. First consider the average relative efficiency of 'all countries'. We observe in two out three time periods a positive productivity change (i.e. $M > 1$) which can be attributed to a combined positive frontier shift and catch-up effect. This indicates that both the best practice and non-best practice observations improve their performances.

More interesting than the average scores of all countries are the country-specific performances. First consider the Walloon utilities which have positive productivity growth in all three periods. Especially the utilities which are lagging behind (the catch-up effect) are improving their performances. On the contrary, the Flemish utilities do not seem to improve their performances. Their efficiency decreases in the first and the last periods. Nevertheless, the best Flemish utilities seem to keep track and steadily improve their performances (i.e. the frontier shift is larger than 1 in each of the three periods). Finally, consider the Dutch utilities where performance improved in the first and last periods thanks to a combination of best practice improvements and catching-up effect. Only in the 2001-2003 period, did the performance decline due to both best practice and non-best practice deterioration.

[6] Notice that the analysis is performed on one aggregate sample consisting of utilities from Wallonia, Flanders and the Netherlands. However, separating the Flemish and Walloon utilities and the Flemish, Walloon and Dutch utilities gave very similar outcomes, as there is a significant correlation among the subsamples of, respectively, more than 0.93 and 0.91.

7. GAMING THE STAKEHOLDERS

The absolute values of the Malmquist index shed some light on the sectors and especially the different directions of the productivity growth provide an interesting observation. The Flemish drinking water sector, which does not undertake a systematic benchmarking, decreases its productivity in two out three time periods. Conversely, both the Walloon and the Dutch drinking water sector, which use benchmarking tools, experience a productivity increase. Nevertheless, the self-regulatory aspect in the Dutch sector seems to take its toll as the productivity change is barely larger than 1 in the last period (the middle period shows even a negative productivity change).

The different directions of the productivity growth could originate from attempts to ‘game’ the regulator (Jamasb *et al.*, 2003, 2004). The three sectors are characterized by the presence of professional associations with strong powers over the sector and over the politicians. As a rule, professional associations do not like regulators or benchmarking. On the one hand, regulation interferes with their activity, and on the other hand benchmarking provides transparency to the sector and distinguishes the best practices from the worst. As the benchmarks generally correspond to a small number of players, the majority of the utilities are embarrassed. Besides, the increased transparency will reduce profits and existing privileges.

In the Netherlands, since the 1990s there has been a huge discussion about the privatization and the creation of water regulators (see also Section 2 and De Witte and Saal, 2008). Influenced by the good outcomes of the energy regulators, governments and think-tankers launched this issue to the public opinion. The water industry heavily rejected the idea arguing that it already presented good performance, even with public ownership and without strict (economic) regulation. Nevertheless, the water utilities committed themselves to further improve their performance by a sunshine regulation (since 1997 and published every three years since 1999). The increased transparency significantly improved the productivity in the period of 1999-2001 (which is confirmed by previous studies: De Witte and Dijkgraaf, 2007; De Witte and Saal, 2008). However, as new debates on the regulatory system have emerged since 2001 (in particular, the government wanted to establish an independent regulator which would apply yardstick competition), the utilities feared its impact and no longer participate in the benchmark (the participation rate to the voluntary benchmark (edition 2003) significantly decreased). As a result, productivity declined (as also observed by De Witte and Saal (2008) by using a profit decomposition). In a third period (2003-2005), the sector was enrolled in new discussions about the ownership of the sector (the Eigendomswet). The water companies, trying to renegotiate with the government their *status quo*, improved their effectiveness and efficiency. The debate effectively triggered productivity again. Notice that in the Netherlands, during this period, the profits were high and above the average of other industries. The monopoly power, even with the benchmarking tool implemented, was kept. Looking from a different angle, benchmarking without practical effects (the ‘sticks’ or the ‘carrots’) always produces limited outcomes. In the first time period, trying to keep the sector as a public domain worked as an effective ‘carrot’. In the second time period, the ‘stick’ of the independent regulator had a perverse effect. Finally, in the third period, the ownership debate worked as a ‘carrot’.

Similar ‘gaming’ can be observed in Flanders. Initially (1999-2001), there was no effective threat to the water utilities so that the quiet life could prevail. This naturally results in a negative productivity growth. After a public debate in the beginning of 2000, privatization and strict regulation were heavily discussed. The water industry reacted promptly and the utilities became more efficient and innovative such that productivity increased. Nevertheless, the Flemish government enacted, in 2003, a law which included an independent regulator. However, as this regulator after 5 years is still not in place and as there are no signs that its establishment is for the near future, the water utilities returned to the quiet life. They acknowledged the inability of the Flemish government in the regulation of the water sector.[7]

In the period under study, the Walloon utilities increased their productivity (although potentially the scope for improvement was larger as well). Similarly to other sectors, there was a debate about privatization and regulation. Even though the utilities tried to avoid the regulatory burden, the Walloon government effectively established a regulator. In spite of applying only the light-handed sunshine approach, there were good outcomes. Besides the benefits of this light-handed regulatory method, the possibility of a change to a more coercive and intervening role in the regulation of the water sector represented also a good incentive for the water industry to go on improving the productivity. The empirical results stress this gaming.

Summarizing, the importance of the water associations in gaming the stakeholders in the water sector in the Low Countries is well highlighted. We observed that the presence of benchmarking or even just the threat of its implementation led to a productivity improvement. Besides, the creation of a regulator can push the water utilities towards the best practices. Furthermore, the Malmquist decomposition reveals that in the Low Countries the productivity change is almost always driven thanks to the catch-up factor and that the change in the technology (frontier shift) is only limited. Indeed, in contrast to other network sectors, technological improvements in the water sector are often negative (due to quality improvements) and very small.

8. CONCLUSIONS

This paper discussed the use of benchmarking in water utilities. Comparing three well comparable regions in Low Countries (well comparable in terms of very similar exogenous factors) by the use of a dynamic non-parametric Malmquist decomposition, we find that the utilities are ‘gaming’ the different stakeholders. The presence and the threat arising from benchmarking and even from regulation create different behaviors in terms of productivity improvement. We observed that if the water utilities are provided with the appropriate incentives they are able to improve their productivity significantly, but if there are no “carrots” (nor “sticks”) the water industry will game the stakeholders abusing from the quiet life and the monopoly power.

[7] Notice that the Flemish region attaches a significant importance to social issues, for example by the provision of free water to householders (until 15 m³). This could be a reason why economic regulation is more difficult to implement.

REFERENCES

Alegre, H. (2007). *Water and wastewater infrastructure asset management*. Research Program, LNEC, Lisbon (in Portuguese).

Alegre, H., J. Baptista, E. Cabrera, J. Cubillo, P. Duarte, W. Hirner, W. Merkel and R. Parena (2006). *Performance indicators for water supply services*. Second Edition, London: IWA Publishing.

Aquawal (2006). *Etude relative à l'élaboration d'indicateurs de performance des services de distribution de l'eau*. Rapport Final. p. 96.

Aubin, D. and F. Varane (2007). Country report Belgium. In *Water and Liberalisation, European Water Scenarios*. Editors Finger, M., J. Allouche and P. Luis-Manso, IWA publishing.

Belgaqua (2007). *Overzicht van de beheersstructuren in België.* Belgische federatie voor de *watersector, p. 32.*

Camp, R. (1989). *Benchmarking, the search for industry best practices that lead to superior performance.* Portland: American Society for Quality Control (ASQC) Quality Press.

Caves, D.W., L.R. Christensen and W.E. Diewert (1982). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50(6), 1393–1414.

Cazals, C., J.P. Florens and L. Simar (2002). Nonparametric frontier estimation: a robust approach, *Journal of econometrics* 106(1), 1-25.

Charnes, A., W.W. Cooper and E. Rhodes (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research* 2(6), 429–444.

Daraio C. and L. Simar (2007). *Advanced robust and nonparametric methods in efficiency analysis*. Series: Studies in Productivity and Efficiency, Springer.

De Witte, K. and E. Dijkgraaf (2007). Mean and bold? On separating merger economies from structural efficiency gains in the water sector. *Tinbergen Institute Discussion Paper* 92/03. Accepted for publication in *Journal of the Operational Research Society*.

De Witte, K. and R. Marques (2007a). Detecting influential observations in frontier models, a robust approach. *Mimeo.*

De Witte, K. and R. Marques (2007b). Designing incentives to local public utilities, an international comparison to the drinking water sector. *CES Discussion Paper Series* DPS 07.32.

De Witte, K. and D. Saal (2008). The regulator's fault? On the effects of regulatory changes on profits, productivity and prices in the Dutch drinking water sector. *CES Discussion Paper Series* DPS 08.28.

Dijkgraaf, E., R. de Jonge, v.d. Mortel, A. Nentjes, M. Varkevisser and D. Wiersma (1997). *Mogelijkheden tot marktwerking in de Nederlandse watersector*. Ministry of Economic Affairs, The Hague.

Dijkgraaf, E., S.A. van der Geest and M. Varkevisser (2007). The efficiency gains of benchmarking Dutch water companies. *Working Paper Erasmus University Rotterdam.*

Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series B* 120(3), 253-281.

Freeman, N., B. Liner and L. Blankenship (1998). Benchmarking. The first step toward improving operating performance. *Water Environment & Technology*, 10(7), 45-48.

Fried, H., C.A.K. Lovell and S. Schmidt (2008). *The measurement of productive efficiency and productivity growth*. Oxford University Press.

Jamasb, T., P. Nillesen and M. Pollitt (2003). Gaming the regulator: a survey. *The Electricity Journal* 16(10), 68-80.

Jamasb, T., P. Nillesen and M. Pollitt (2004). Strategic behaviour under regulatory benchmarking. *Energy Economics* 26(5), 825-843.

Kingdom, B., J. Knapp, P. LaChange, and M. Olstein (1996). *Performance benchmarking for water utilities*. Denver: AWWARF and AWWA.

Kingdom, B. (1998). Use of performance indicators and performance benchmarking in the North American water industry—findings from studies recently completed for AWWA and WEF research foundations. *Aqua*, 47(6), 266–296.

Lafferty, A. and L. William (2005). Benchmarking performance indicators for water and wastewater utilities. AWWARF, Denver.

Marques, R. (2005). *Regulation of public services*. Sílabo Press, Lisbon.

Marques, R. (2006). A yardstick competition model for Portuguese water and sewerage services regulation. *Utilities Policy* 14(3), 175–184.

Marques, R. and K. De Witte (2008). Towards a benchmarking paradigm in the European public water and sewerage services. Forthcoming *Public Money & Management*.

Parena, R., E. Smeets and I. Troquet (2002). *Process benchmarking in the water industry*. London: IWA Publishing.

SERV (2007). *Informatiedossier. De Watersector: analyse en uitdagingen*. Sociaal Ecomische Raad van Vlaanderen, p. 141.

Van Dijk, P., M. Schouten, K. Swami and M. Kooij (2007). *Country report on the Netherlands. In Water and Liberalisation, European Water Scenarios*. Editors Finger, M., J. Allouche and P. Luis-Manso, IWA publishing.

Vewin (2007a). *Reflections on performance 2006*, Association of Dutch Water Companies, Rijswijk.

Vewin (2007b). Internationale benchmark krijgt vervolg. *Waterspiegel* 10(1), 19-21.

Table 3 - Summary statistics

Wallonia		**Length mains**	**Staff**	**Connect.**	**Distributed volume**	**Flanders**		**Length mains**	**Staff**	**Connect.**	**Distributed volume**
1999	mean	1,362	124	38,489	4,552,253	1999	mean	9,941	608	363,370	60,426,562
	st dev	5,362	360	127,508	14,573,583		st dev	11,229	547	378,537	45,051,837
	max	28,000	1,372	639,526	73,041,454		max	28,588	1,506	987,768	112,205,738
	min	23	1	702	59,009		min	943	46	27,098	4,593,270
2001	mean	1,151	130	40,291	4,657,146	2001	mean	10,192	615	394,575	60,841,608
	st dev	4,099	373	128,878	14,614,026		st dev	11,279	536	381,398	44,355,566
	max	21,390	1,464	642,213	72,478,755		max	28,844	1,494	1,012,050	113,582,574
	min	20	1	700	62,320		min	953	48	27,559	4,385,903
2003	mean	1,278	133	41,094	4,734,663	2003	mean	10,422	614	416,254	63,692,763
	st dev	4,330	384	132,712	15,055,725		st dev	11,354	527	388,923	41,801,151
	max	22,547	1,447	664,466	74,828,498		max	29,161	1,479	1,032,816	117,001,068
	min	25	2	705	63,754		min	997	49	27,844	4,831,999
2005	mean	1,300	136	43,116	7,080,212	2005	mean	10,358	628	420,957	77,712,622
	st dev	4,500	391	141,534	23,264,967		st dev	11,682	540	396,652	53,213,663
	max	23,432	1,470	713,248	118,428,033		max	29,509	1,517	1,052,128	147,381,406
	min	25	2	704	98,720		min	101	46	28,015	5,458,590
Netherl.		Length mains	Staff	Connect.	Distributed volume	Netherl.		Length mains	Staff	Connect.	Distributed volume
1999	mean	4,993	365	280,871	56,186,100	2003	mean	5,514	279	308,223	60,035,719
	st dev	2,939	257	155,012	35,341,380		st dev	3,196	181	165,476	36,096,447
	max	10,309	980	620,608	145,112,000		max	11,783	593	649,192	153,143,389
	min	680	61	83,003	11,877,650		min	683	50	87,127	12,278,332
2001	mean	5,138	333	287,887	56,200,050	2005	mean	5,650	260	314,404	58,540,704
	st dev	2,989	229	158,818	34,425,614		st dev	3,279	167	168,566	37,438,291
	max	10,423	844	639,000	139,325,000		max	11,981	529	681,812	165,710,720
	min	696	59	85,454	12,025,097		min	705	47	87,127	11,546,090

Table 4 - Results Malmquist decomposition

	1999-2001			2001-2003			2003-2005		
All countries	Malmquist	Catch-up	Frontier shift	Malmquist	Catch-up	Frontier shift	Malmquist	Catch-up	Frontier shift
Average	1.0827	1.0779	1.0082	0.9918	0.9985	0.9954	1.0281	1.0325	1.0029
St. Deviation	0.3298	0.3306	0.0627	0.2149	0.2226	0.0645	0.1728	0.1987	0.0775
Minimum	0.6295	0.5486	0.8901	0.6211	0.6249	0.8795	0.4819	0.4927	0.8354
Maximum	2.8420	2.7908	1.1474	1.9128	2.0425	1.1538	1.6247	1.7263	1.2318
Wallonia									
Average	1.1180	1.1124	1.0084	1.0273	1.0309	0.9993	1.0753	1.0854	0.9966
St. Deviation	0.4484	0.4459	0.0590	0.1555	0.1596	0.0720	0.1607	0.1907	0.0653
Minimum	0.6295	0.5486	0.8901	0.7459	0.7175	0.8795	0.8283	0.7860	0.8654
Maximum	2.8420	2.7908	1.1474	1.3189	1.3636	1.1538	1.6247	1.7263	1.1622
Flanders									
Average	0.9770	0.9489	1.0330	1.0364	1.0349	1.0014	0.8911	0.8796	1.0136
St. Deviation	0.0758	0.0845	0.0804	0.1185	0.1017	0.0520	0.2373	0.2242	0.1062
Minimum	0.8844	0.8529	0.9165	0.9187	0.9538	0.9109	0.4819	0.4927	0.8880
Maximum	1.0883	1.0528	1.1143	1.2322	1.2058	1.0350	1.0870	1.0625	1.1810
The Netherlands									
Average	1.0616	1.0635	1.0017	0.9326	0.9458	0.9887	1.0010	1.0019	1.0085
St. Deviation	0.1025	0.1160	0.0649	0.2863	0.3021	0.0587	0.1548	0.1858	0.0877
Minimum	0.8733	0.8210	0.8976	0.6211	0.6249	0.8832	0.6728	0.5462	0.8354
Maximum	1.1992	1.2195	1.0945	1.9128	2.0425	1.0812	1.2970	1.3838	1.2318